Nucl ic acid molecules encoding pr teins which influ nce bone devel pment

Ins a'

The present invention relates to nucleic acid molecules encoding proteins which influence the bone development of mammals, the encoded proteins, and diagnostic and pharmaceutical compositions containing such nucleic acid molecules or proteins. Moreover, the invention relates to transgenic non-human mammals which are transformed by the herein-described nucleic acid molecules or which show a modified expression of the herein-described proteins.

In humans, a number of hereditary diseases resulting in impaired growth and development of the bones are known. These, for instance, include spondyloepiphyseal dysplasias and achondroplasia. The exact genetic factors causing such disorders are, as a rule, unknown and therapeutic approaches or diagnostic methods for an early detection are in most cases not available.
The elucidation of the factors causing such growth and development disturbances and the provision of possible therapeutical approaches and diagnostic methods for an early detection of such disturbances require the identification and isolation of genes participating in the regulation of corresponding growth and development processes.

Hence, the technical problem underlying the present invention is the provision of nucleic acid molecules, the expression product of which influences growth and development processes, in particular relative to bones, in animals and humans.

This problem is solved by the provision of the embodiments as characterized in the claims.

Thus, the present invention relates to nucleic acid molecules comprising a nucleotide sequence encoding the amino acid sequence depicted in SEQ ID No. 9 or in SEQ ID No. 14, and nucleic acid molecules comprising the nucleotide sequence depicted in SEQ ID No. 8 or SEQ ID No. 13, and in particular comprising the coding region. Such nucleic acid molecules can contain the corresponding

coding regions in a continuous form or in a form interrupted by non-coding regions. Consequently, such molecules can also be genomic sequences, in which the coding regions (exons) are interrupted by non-coding regions (introns). Surprisingly, the protein encoded by such a nucleic acid molecule has been found to be a protein, the inactivation of which in mammals has the effect that the bones, except for the scull bones, become longer. Such nucleic acid molecules were found in connection with the production of a so-called transgenic "donor" mouse, that is to say a mouse which was to serve as a donor of an artificial protein. This artificial protein was to be expressed in particular tissues of the "donor" mouse, without, however having any function in this mouse. The protein should become effective only after crossbreeding the donor mouse with a suitable transgenic recipient mouse and should activate particular genes of the recipient mouse. Transgenic donor mice have already been produced from time to time. Normally, they do not show a phenotype, because the artificial gene is simply injected into fertilized egg cells and integrates into any one region of the murine genome on a purely random basis. As only about 5% of the genome are coding regions, the probability that a defect is caused in an essential gene is relatively small. Moreover, the mammal genome is diploid, that is to say, all genes are present in duplicate. Hence, most mutations are recessive, that is to say they do not show up: the mutated gene has a fully functioning copy as a counterpart, which is able to compensate for the defect generated.

Surprisingly, the donor mouse produced shows an extremely conspicuous phenotype: all bones (except for the scull) are 1.3 to 1.5 times longer. As a consequence, the transgenic mouse is about 1.5 times longer than the corresponding wildtype (see Fig. 1). This phenotype is dominant and is stably passed on, that is to say in crossbreeding a transgenic mutant with a healthy wildtype mouse, 50% of the offspring show the above-described phenotype.

Genetic analysis of this mouse showed that a gene was inactivated by the insertion into the genome of the DNA for the artificial protein to be produced in the mouse. In order to find out which gene (or which genes) is/are responsible for the phenotype found, the mutated region of the genome of the transgenic mouse was subcloned in bacteria. The localization of the mutated region in the genome of the mouse and the subsequent subcloning were possible because the nucleotide sequence of the

inserted artificial gene was known, and this information could be utilized in corresponding molecular biological experiments.

For identifying the gene, hereinafter called LOBO-gene ("long bones"), 6 kb of the subcloned region of the transgenic mouse were sequenced and first 87 kb (SEQ ID Nos. 5 and 6) and then altogether 138 kb (SEQ ID Nos. 10 to 12) were sequenced of the corresponding homologous region of the wildtype mouse. A detailed computer analysis of the sequence data led to the identification of a gene which consists of at least 13 coding segments ("exons") and is at least 110 000 bases long, but probably much longer. The first identified coding region of the murine genomic sequence carries the information for 393 amino acids (see SEQ ID No. 2). On the basis of the murine sequence data obtained, a DNA probe was constructed, which was used to isolate a human P1 clone carrying the human LOBO homologous gene. The sequence of the first sequenced 13.3 kb long region is depicted in SEQ ID No. 7. The sequence of the isolated and identified coding regions (exons) of this gene is depicted in SEQ ID No. 3 as is the amino acid sequence derived therefrom. The sequence of the subsequently sequenced 311 kb long region is depicted in SEQ ID Nos. 15 to 21. The sequence of the coding regions identified therein (exons) is depicted in SEQ ID No. 13, the amino acid sequence derived therefrom in SEQ ID No. 14. Using the genomic sequence information, it was subsequently possible to isolate a complete 3100 bp long cDNA of the murine LOBO gene (SEQ ID No. 8). Of these 3100 bp 1857 bases from the 3'-end have been also elucidated by the genomic sequencing. Hence, the exon/intron structure is known for this section: there are 12 exons, enumerated from the 3'-end in increasingly higher figures, that is to say the exon positioned at the most proximate 3'- end is numbered 1, the outermost exon identified so far is numbered 12. By means of the sequence data provided by the present invention, it is possible to isolate and characterize the still missing regions of the gene by standard methods, for instance chromosomal walking. The murine cDNA carries the information for a protein having a length of 870 amino acids (SEQ ID No. 9). A sequence comparison between the amino acid sequence derived from the murine cDNA and the known sequences showed that the encoded protein has a certain homology to a protein of C. elegans (data base

accession No. Q09568), and homologies to the Dis3-protein family and RNAseII protein family.

From the above it follows that the nucleic acid molecules of the invention encode a protein, the modification of which, in particular the reduction and/or inactivation in animals, preferably in vertebrate, preferably in mammals and more preferably in mice results in an elongation of the bones except for the scull bones. An elongation, in this connection preferably means an elongation by a factor of at least 1.2, preferably by a factor of 1.3, and more preferably by a factor in the range of 1.3 to 1.5.

As used herein, the term "modification", in particular reduction and/or inactivation, may comprise quantitative and/or qualitative deviations.

Thus, on the one hand, from a quantitative point of view, the term "modification", in particular reduction and/or "inactivation", means that the expression of the protein is reduced, preferably by at least 50%, compared to the wildtype, and is more preferably repressed altogether. The analysis of the mutation in the genome of the above-described donor mouse showed that the insertion of the artificial gene is located within an intron of the LOBO gene and has led to the deletion of 11 base pairs. The latter should not pose a problem in the intron, as this area is not a coding region anyway. Hence, it can be assumed that the artificial DNA insertion leads to a disorder in the maturation ("splicing") of the mRNA, as the artificially inserted gene contains splicing signals. This presumably leads to a so-called "aberrant splicing". In consequence, a functioning mRNA is prevented from being formed and the corresponding protein cannot be produced. In actual fact, the experimental investigation of the LOBO expression (by "Northern blot") has shown that heterozygous LOBO mice produce only about half the amount of mRNA produced by the wildtype mouse. In homozygous LOBO mice no LOBO mRNA whatsoever can be detected in Northern blot. Hence, it can be assumed that the mutation in the transgenic LOBO mouse switches off gene expression on the post transcriptional level. Apparently, the amount of LOBO protein produced in the heterozygous mice then already falls below a critical threshold value, which then leads to the dominant phenotype found.

Hence, within the present invention, the t rm "modification", in particular reduction and/or "inactivation" preferably means that the amount of transcripts encoding the protein described, is reduced in the cells compared to cells of corresponding wildtype animals by at least 50%, preferably by at least 70%, more preferably by at least 90%. In an especially preferred embodiment "modification", in particular reduction and/or inactivation, means that no transcripts encoding the protein described herein can be detected any more. The amount of transcripts can be detected by techniques known to a skilled person, for instance by Northern blot analysis.

On the other hand, from a qualitative point of view, the term "modification", in particular reduction and/or inactivation, means that a LOBO protein modified in the amino acid sequence is expressed, in particular a protein which has completely or largely lost its biological function. Such proteins can be shortened forms, forms, which show deletions or insertions, forms which have one or more point mutations or forms which are combinations of one or more forms of this modification. For instance, as the above-described transgene-insertion in the transgenic LOBO mouse does not affect the expression signals (promoter, enhancer etc.), it could be assumed that at least a shortened and in addition chimeric LOBO mRNA is produced from the native transcription start to the splice signal in the inserted sequence. However, a polyadenylation signal is missing from the transgene-insertion, which leads to a non-polyadenylated RNA. This RNA should possess a distinctly reduced stability vis-à-vis the normal LOBO mRNA. That is to say, the amount of this chimeric RNA should be relatively small and below the Northern blot detection limit. In fact, this chimeric RNA has not been detected in Northern blot so far. However, the much more sensitive RT-PCR method made it possible to verify the existence of this postulated chimeric RNA. Hence, this RNA can be assumed to cause the formation of a shortened LOBO protein, which carries some amino acids from the artificial gene at its COOH end.

Hence, there may be two causal factors for the long bone phenotype: (a) the amount of transcripts encoding the complete LOBO protein falls below the critical

value*) because of the transgene-insertion (loss of function mutation) and/or (b) a shortened, chimeric LOBO protein is produced which shows only partial functions of the LOBO protein or modified functions compared to the LOBO protein (gain of function mutation).

Moreover, the "modification", in particular the reduction and/or inactivation, of the protein encoded by the nucleic acid molecules of the invention, preferably leads to at least one of the following modifications in mice:

(a) The bones show significantly thickened growth zones from a histological point of view (see Figure 4). Preferably, this stems from a marked increase in the number of cells in the growth zone (chondrocytes). Moreover, these chondrocytes are distinctly larger than those of corresponding wildtype mice;
(b) life expectancy is dramatically shortened, it is 40 weeks as a maximum and about 25 weeks on the average (in wildtype mice, the mean life expectancy is 1 to 2 years).

The amino acid sequences of the murine and human proteins encoded by the nucleic acid molecules of the invention were compared with those of known proteins. The comparison showed that the amino acid sequence possesses regions highly conserved between organisms ranging from mammals (humans, mice) to invertebrates (C. elegans), unicellular eukaryotes (Saccharomyes cerevisiae, Schizosaccharomyces pombe) and prokaryotes (Leuconostoc). A relationship analysis showed in particular that the murine and human LOBO proteins constitute a group of their own (see Figure 6) which is, however, related to two other protein groups. The VacB- and the RNAse-type-II-proteins from bacteria constitute one group. The Dis3-homologous proteins from different eukaryotes, ranging from mammals to unicellular yeasts constitute a second group.
Because of the clear relationship to the two afore-mentioned groups of proteins, the function of the proteins encoded by the nucleic acid molecules of the invention can be estimated. It is assumed that because of their structural similarity to the afore-

*) Translator's note: "Should read threshold value"

mentioned two other protein groups, these proteins also have similar functions. The following functions of the LOBO proteins can be postulated on this basis:

(a) they play an important role in the regulation of the cell cycle (mitosis control) (proven for Dis3 from S. pombe; here the loss of function of the gene results in the loss of the capability of the cells to divide);

(b) because of their bearing on the cell cycle control, the conclusion suggests itself that the LOBO proteins might also play a part in carcinogenesis (so far, this has been proven for Dis3 from Homo sapiens; the results shown in Figure 5 obtained in a Northern blot analysis with a LOBO probe and RNA from diverse tumor tissues support this);

(c) the LOBO protein is most probably able to bind RNA (proven so far for the LOBO-type SSDI protein from S. cerevisiae and for the VacB- and RNAse type II proteins); and/or

(d) the LOBO protein has at least one protein binding partner. This is presumably a G-protein or a G-protein-controlling protein (proven for Dis3 from S. pombe, which binds to the G-protein regulator RCC1 and controls its activity).

Because of the impressing bone phenotype and because of the relationship to the Dis3-protein family, the provision of the nucleic acid molecules of the invention is of great importance both from a scientific and a clinical point of view. On the one hand, its further investigation could help understand the cell cycle control still better. This is in particular important in cancer research. On the other hand, the nucleic acid molecules of the invention could be responsible for human growth disorders, not caused by nutrition or hormones.

The present invention also relates to nucleic acid molecules, the complementary strand of which hybridizes with one of the above-described nucleic acid molecules of the invention and which encode a protein having the above-mentioned properties.

The term "hybridization" as used herein means hybridization under conventional hybridization conditions, preferably under stringent conditions as for instance described in Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd edition, (1989), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). In this context the term "stringent conditions" means that hybridization only occurs if the sequence identity is at least 90%, preferably at least 95% and more preferably of at least 97% over the entire length of the molecule hybridizing to the molecule of the invention. Specific examples of stringent and non-stringent hybridization conditions are published for instance in Hames and Higgins (editors), "Nucleic acid hybridization: A practical approach", IRL press, Oxford-Washington DC, 1985. An example of stringent hybridization conditions is, for instance, filter hybridization to polynucleotide probes, wherein the filter is washed in 0.1 x SET buffer and 0.1% SDS solution for 20 minutes at 68°C. An example of non-stringent hybridization conditions is for instance filter hybridization with polynucleotide probes, wherein the filter is washed in 2 x SET buffer and 0.1% of SDS solution for 20 minutes at 50°C. Nucleic acid molecules which hybridize to the nucleic acid molecules of the invention can, in principle, be derived from any animal organism which expresses such a protein. Molecules encoding corresponding proteins from higher animal organisms are preferred, and they preferably originate from vertebrates, and more preferably from mammals and in particular from mice or humans.

Nucleic acid molecules which hybridize with the molecules of the invention can, for instance, be isolated from genomic or cDNA libraries. Such nucleic acid molecules can be identified and isolated with the use of the nucleic acid molecules of the invention or parts of these molecules or reverse complements of these molecules, for instance by hybridization according to standard methods (see for instance Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, second edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) or amplification by PCR.

For instance, nucleic acid molecules, which have exactly or substantially the nucleotide sequence which is indicated in SEQ ID No. 8 or 13 or comprise parts thereof can be used as hybridization probes. The fragments used as hybridization probe can also be synthetic fragments which are prepared by conventional

synthesis techniques and the sequence of which is substantially identical to that of a nucleic acid molecule of the invention. Once genes have been identified and isolated which hybridize to the nucleic acid sequences of the invention, the sequence should be determined and the properties of the proteins encoded by this sequence should be analyzed.

The molecules hybridizing to the nucleic acid molecules of the invention in particular comprise fragments, derivatives and allelic variants of the above-described nucleic acid molecules encoding a protein possessing the above-described properties. In the present context, the term "derivative" means that the sequences of these molecules differ from the sequences of the above-described nucleic acid molecules at one or more positions and have a high degree of homology to these sequences. In this connection, homology means a sequence identity on the amino acid level over the entire length of at least 70%, in particular an identity of at least 80%, preferably more than 90%, especially preferably more than 95%, and in particular of at least 97%. Moreover, homology preferably means a sequence identity of at least 60 %, preferably at least 70%, more preferably at least 85% and most preferably of at least 95% on the nucleic acid sequence level. Deviations from the above-described nucleic acid molecules can, for instance, be caused by deletion, addition, substitution, insertion or recombination.

Moreover, homology means that there exists functional and/or structural equivalence between the corresponding nucleic acid molecules or the proteins encoded by them. The nucleic acid molecules which are homologous to the above-described molecules and are derivatives of these molecules are, as a rule, variations of these molecules representing modifications which have the same biological function. The variations can be naturally occurring ones, for instance sequences from other animal species or mutations, and said mutations may have occurred naturally or may have been introduced by specific mutagenesis. Moreover, the variations may be synthetically prepared sequences. The allelic variants can be both naturally occurring variants and variants prepared synthetically or by recombinant DNA techniques.

The proteins encoded by different variants of the nucleic acid molecules of the invention possess certain characteristics they have in common. These may for

instance include biological activity, molecular weight, immunological reactivity, conformation etc., and physical properties, such as for instance mobility in gel electrophoresis, chromatographic behavior, sedimentation coefficients, solubility, spectroscopic properties, stability, pH optimum, temperature optimum, etc.

The proteins encoded by the nucleic acid molecules of the invention preferably have the same biological function or activity as that described above for the murine protein, i.e. in the case of a modification, in particular reduction and/or inactivation of these proteins, vertebrates can show the above-described disturbances in bone development.

Sub B1 It is particularly preferred for the protein encoded by the nucleic acid molecule of the invention to comprise at least one of the following two consensus sequences.

Consensus 1:

EFMLLANXXVAXXIXXXFPXXALLRRHXXP

Consensus 2:

HZALNVXXZTHFTSPIRRZXDVIVHRLLAAALGY

Moreover, the present invention relates to nucleic acid molecules, the sequence of which deviates from the sequence of one of the above-described nucleic acid molecules because of the degeneracy of the genetic code.

The nucleic acid molecules may be any nucleic acid molecules, in particular DNA or RNA molecules, for instance cDNA, genomic DNA, mRNA etc. They may be naturally occurring molecules, or molecules prepared by genetic engineering or chemical synthetic methods.

Examples of genomic murine or human sequences are given in SEQ ID Nos. 5, 6, 7, 10 to 12 and 15 to 21. The murine gene was localized in band 1D on murine chromosome 1, using "fluorescent in situ hybridization" (Fish) on whole murine metaphase chromosomes. This band is synthenic to band 2q35, in particular to region 2q35-37 on human chromosome 2. This segment also contains a gene for alkaline phosphatase, the exact position of which is known in the literature. The analysis of the murine and human genomic sequences carrying a nucleic acid molecule of the invention showed that in both cases the gene for the alkaline

phosphatase is located about 20 kb downstream of the LOBO gene, with the result that the chromosomal localization of the latter can be very precisely specified. With the help of the nucleic acid molecules disclosed in the present invention, it is possible for a skilled person to isolate homologous sequences from other organisms, in particular mammals, by means of known techniques.

Moreover, the invention relates to vectors, in particular plasmids, cosmids, viruses, bacteriophages and other vectors commonly used in genetic engineering which contain the above-described nucleic acid molecules of the invention. These are, preferably, vectors which are suitable for gene therapy.

In a preferred embodiment, the nucleic acid molecules contained in the vectors are linked to regulatory elements ensuring the expression in prokaryotic or eukaryotic cells. In this context, the term "expression" can mean both transcription as well as transcription and translation. Here, regulatory elements in particular include promoters. The number of promoters available for the expression of a nucleic acid molecule of the invention in prokaryotic cells include for instance the E. coli lac- or trp-promoter, the P_R- or P_L- promoter of the λ phage, lacI, lacZ, T3, T7, gpt etc. Eukaryotic promoters are, for instance, the CMV immediate early promoter, the HSV promoter, the thymidin kinase promoter, the SV40 promoter, LTRs of retroviruses and the mouse metallothionin-promoter. A great number of expression vectors for the expression in prokaryotic or eukaryotic cells have been described, for instance for eukaryotes pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) or GEM1 (Promega Biotec, Madison, Wi, USA), pSV2CAT, pOG44 and for prokaryotes pQE70, pQE60, pBluescript SK, etc. Vectors of the invention may contain not only promoters but also elements to increase transcription further, such as for instance the so-called transcription enhancers. Examples thereof are the SV40 enhancer, the polyoma enhancer, the cytomegalovirus early promoter enhancer and adenovirus enhancer.

The present invention also relates to host cells, in particular prokaryotic or eukaryotic host cells, which are transformed with a nucleic acid molecule or a vector

of the invention. Examples of such cells are bacterial cells, such as for instance E. coli, Streptomyces, Bacillus, Salmonella typhimurium; fungal cells, such as yeast cells, in particular Saccharomyces cerevisiae; insect cells, such as Drosophila or SF9 cells; animal cells, such as CHO or COS cells, plant cells etc.

Moreover, the present invention relates to a method for producing a protein encoded by a nucleic acid molecule of the invention, wherein a host cell according to the invention is cultured under conditions permitting the expression of the protein, and the protein is subsequently recovered from the cells and/or the culture medium. Methods for the expression of foreign proteins in different species of host cells and for recovering the protein produced are known to a skilled person.

Moreover, the invention relates to a protein which is encoded by a nucleic acid molecule of the invention or is obtainable by the method of the invention.

Moreover, the present invention relates to antibodies, directed against the proteins of the invention. Preferably, such antibodies specifically recognize a protein of the invention, that is to say they do not show any substantial cross reaction with other proteins. In this connection, the term "antibody" comprises both monoclonal and polyclonal antibodies, as well as the fragments of antibodies, for instance Fab fragments, said fragments recognizing a protein of the invention. The term "antibody" also comprises chimeric antibodies and humanized antibodies. Methods for producing monoclonal or polyclonal antibodies are known to a skilled person and have been described. Monoclonal antibodies can be prepared for instance by the hybridoma technique (Köhler and Milstein, Nature 256 (1975), 495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., Immunology Today 4 (1983), 72) or the EBV-hybridoma technique (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Lise, Inc. (1985), 77-96).

Moreover, the present invention relates to nucleic acid molecules having a length of at least 15, preferably more than 50 and particularly preferably more than 200 nucleotid s which specifically hybridize to a strand of a nucleic acid molecule of the

invention. As used herein, "specifically hybridize" means that these molecules hybridize to nucleic acid molecules encoding a protein of the invention, but do not hybridize to nucleic acid molecules encoding other proteins. In this connection, hybridizing preferably means hybridizing under stringent conditions (see above). Such nucleic acid molecules can, for instance, be used as primers for PCR amplification or as hybridization probes. The invention in particular relates to the nucleic acid molecules which hybridize with transcripts of nucleic acid molecules of the invention and can thereby prevent their translation. Such nucleic acid molecules can, for instance, be components of antisense constructs or ribozymes.

Moreover, the present invention relates to diagnostic compositions containing a nucleic acid molecule or a vector, a protein and/or an antibody according to the invention. The nucleic acid molecules of the invention can, for instance, be used to determine the localization of the corresponding gene on a chromosome. This can elucidate the correlation to genes associated with particular diseases. A method for determining the localization is for instance "fluorescent in-situ hybridization" (Fish) which is described in Verma et al. (Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York (1988)). Moreover, the nucleic acid molecules of the invention can be used to determine whether particular individuals have mutations in the corresponding sequences. Similarly, antibodies can be used as reagents to detect the presence of a protein of the invention in a sample.

The present invention also relates to pharmaceutical compositions containing a nucleic acid molecule, vector, protein and/or antibody according to the invention, optionally in combination with a pharmaceutically acceptable carrier. For instance, nucleic acid molecules or vectors of the invention can be used in gene therapy, in order to treat pathological conditions attributable to a dysfunction of the corresponding gene, for instance to too low or too high an expression of the protein of the invention in an individual. The nucleic acid molecules can in particular be used in connection with gene targeting and/or gene replacement, in order to reconvert a mutated gene into a functional form or in order to generate a mutated gene by homologous recombination (see for instance Mouellic, Proc. Natl. Acad.

Sci. USA 87 (1990), 4712-4716; Joyner, Gene Targeting, A Practical Approach, Oxford University Press). Similarly, a protein or antibody of th invention can be used, in order to possibly control the amount of corresponding protein in an individual.

Examples of suitable pharmaceutically acceptable carriers are known to a skilled person and, for instance, include phosphate-buffered salines, water, emulsions such as oil/water emulsions, sterile solutions etc. Compositions containing such carriers can be formulated according to conventional methods. The pharmaceutical compositions can be administered to the individual in question in a suitable dose. Administration routes are, for instance, the intravenous, intraperitoneal, subcutaneous, intramuscular, topical or intradermal route. Here, dosage depends on many factors, such as the size, sex, weight and age of the patient and the type of the specific compound administered, the manner of administration etc. Generally, the daily dose is 1 μg to 10 mg of units per day. In connection with the intravenous injection of DNA, dosages of 10^6 to 10^{22} copies of the DNA molecule are usual. The compositions can be administered locally or systemically. Generally, administration will be parenteral, for instance intravenous. DNA can also be administered directly at the target site, for instance by biolistic application.

Moreover, the present invention relates to a method for preparing a transgenic, non-human animal, preferably a transgenic mouse, comprising the introduction of a nucleic acid molecule or vector into a germ cell, embryonic cell, egg cell or a cell derived therefrom. The non-human animal used as the donor of the cells in such a method may, for instance, be a healthy, non-transgenic animal or an animal which has a disease or disorder, in particular an animal which suffers from a growth disturbance, preferably a growth disturbance relating to the bones. Such a disease or disorder can be innate or can have occurred naturally or may have been caused by genetic engineering, for instance by the introduction and/or expression of a foreign DNA.

Moreover, the present invention relates to transgenic, non-human animals which are transformed with a nucleic acid molecule or vector of the invention or which are

obtainable by the above-described method. The nucleic acid molecule of the invention is preferably stably integrated in th genome of such transgenic animals. Examples of transgenic animals are transgenic rats, hamsters, dogs, monkeys, rabbits or swine. Transgenic mice are preferred.

The present invention also relates to transgenic non-human animals, in particular mice, in which the expression of the protein of the invention is reduced. Such a reduction can, for instance, be achieved by genetic modification of the cells of the animals, with the result that they express an antisense RNA, a ribozyme or a co-suppression RNA leading to reduced expression of the proteins of the invention in the cells. Alternatively, reduced expression of the proteins of the invention can also be achieved by the inactivation of at least one, preferably all copies of a gene corresponding to a molecule of the invention in the genome of the cells. Such inactivation can, for instance, be achieved by the insertion of foreign DNA into coding or non-coding regions of the corresponding gene. The inactivation of the regulatory regions of the gene is also possible. Moreover, the deletion of regions of the gene is possible.

Furthermore, the present invention also relates to the possibility of activating nucleic acid molecules of the invention in vivo, that is to say in cells, cell cultures or organisms (gene activation). This can, for instance, be achieved by the insertion of a promoter into the genome of a cell containing a nucleic acid molecule of the invention, the promoter being inserted in front of the nucleic acid molecule of the invention. This promoter is, for instance, a constitutive promoter and ensures very high expression or a promoter which is inducible, and when being induced ensures very high expression.

The plasmids HSL1 and HSL2 (HSL = Homo sapiens LOBO) prepared within the scope of the present invention were deposited according to the requirements of the Budapest Treaty at the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) in Braunschweig, Federal Republic of Germany, which is recognized as an

international depositary institution, on March 25, 1998 and March ??, 1999 with the accession numbers DSM 12073 and DSM 12715, respectively.

Figure 1 shows a heterozygous LOBO mouse with an insertion in the LOBO gene (top) compared to a wildtype mouse. The two animals are siblings and are about 6 weeks old.

Figure 2 shows the first pursued sequencing strategy for sequencing the murine and human LOBO gene. As at first only the 3'-end of the gene was sequenced, the exons starting at the 3'-end were numbered 1, 2, 3 etc. Three murine wildtype cosmid clones (middle) were sequenced, two plasmid clones were sequenced from the transgenic LOBO mouse (top) and a human P1-clone (bottom) was sequenced. The arrows denote the exons known for the time being. Seven exons were located on the genomic sequence, the eighth exon at first only existed on an EST clone. The plasmid clones from the transgenic LOBO mouse (top) contain the introduced artificial gene and the adjacent murine sequences. These murine sequences are identical to the corresponding sequences of the wildtype mouse except for 10 base pairs, which have been replaced in the transgenic mouse by the artificial gene.

Figure 3 shows a sequence comparison between the human (HS) and murine (MM) LOBO proteins and the eukaryotic Dis3-homologous and Dis3-type proteins.

Figure 4 shows a histological thin section through a bone growth zone of the LOBO mouse (right-hand side) compared to the wildtype (left-hand side). The exaggerated bone growth of the LOBO mouse is also histologically reflected: compared to the wildtype, the growth zone (proliferative zone) of the LOBO bones is significantly thickened. Moreover, the number of the hypertrophic chondrocytes in the growth zone is distinctly increased.

Furthermore, the chondrocytes of the LOBO mutant are distinctly larger than those of the wildtype mouse.

Figure 5 shows a Northern blot with RNA from human tumor tissues. A commercially available Northern blot (company Clontech) which contains RNA from 8 different human tumor tissues, was hybridized to a radioactively labeled LOBO probe. This probe was prepared by PCR amplification of a human LOBO EST clone. There are significant differences in the expression in the individual tissues: LOBO is overexpressed in chronically myelogenic leukemia (lane 3) and in melanoma (lane 8). In Burkitt lymphoma, by contrast, it does not seem to be expressed at all.

(1) promyelotic leukemia
(2) Hela cell line
(3) chronic myelogenic leukemia
(4) lymphoblastic leukemia
(5) Bukitt lymphoma
(6) colorectal adenocarcinoma
(7) lung cancer
(8) melanoma

Figure 6 shows an analysis of the relationship between LOBO and similar proteins. The analysis was made with the program PHYLIP 3.5 ("Neighbour Joining Method"). As can be seen from the pedigree, the murine and human LOBO proteins represent a group of their own, which is, however, related to the eukaryotic Dis3 proteins and the proteins of the RNAse II-type. Although some of the afore-mentioned invertebrate organisms have been sequenced completely or at least largely, no genuine LOBO homologue has been found among them.

Figure 7 shows an X-ray image of the leg of a LOBO mouse (right-hand side) compared to the wildtype (left-hand side). Every single bone of the LOBO leg is longer by the factor of 1.5 than that of the wildtype.

Figure 8 shows the phenotype of an adult heterozygous LOBO mouse. The incessant bone growth leads to a pronounced deformation of the whole animal, its mobility is highly reduced. Because of the deformation, female LOBO mice can be mated in exceptional cases only, to the effect that homozygous offspring can only be obtained in rare cases. The LOBO males are capable of reproduction.

Figure 9 shows a clone chart and a gene model of the murine LOBO gene on chromosome 1, band D. Seven overlapping cosmid clones were sequenced (A), which result in a continuous genomic sequence of 138,884 base pairs. A sequence comparison with the murine LOBO cDNA allowed 12 LOBO exons to be identified so far (B). Parallel sequencing of the LOBO gene of the transgenic mouse and the wildtype mouse allowed the position of the artificially integrated DNA segment (cassette) to be localized. It is located in the intron between exons 8 and 7.

Figure 10 shows a clone map and a gene model of the human LOBO region on chromosome 2q37. Four overlapping BAC/PAC clones were sequenced (B), which form a continuous genomic sequence of 314,449 base pairs. A sequence comparison with the murine LOBO-cDNA has allowed 11 human LOBO exons to be identified so far (A). Moreover, 6 further genes were identified in the 3' region of the LOBO gene. Five of these genes were known on the cDNA level, the sixth gene is new. Although there exist EST sequences corresponding to this gene in the data base, the localization and genomic structure of this gene have been unknown so far. The chromosomal position of the LOBO gene has been

unambiguously verified by the identification of the STS marker WI 9864 which has been mapped on 8q24.

(1) heat-stable alkaline phosphatase, exons from the data base entry M19159
(2) heat-stable alkaline phosphatase, exons from the data base entry X55958
(3) heat-stable alkaline phosphatase, exons from the data base entry M31008.
(4) unknown gene identified by computer analysis
(5) nicotine-dependent acetyl choline receptor, delta subunit, exons from the data base entry X55019
(6) nicotine-dependent acetyl choline receptor, gamma subunit; exons from the data base entry X55019

The following examples illustrate the invention.

Example 1

Detection of a mouse showing modified bone growth

In connection with the investigation of a particular artificial protein, a transgenic mouse was produced, which was to serve as a donor mouse, i.e. as a donor of the artificial protein. This protein was to be expressed in particular tissues of the donor mouse, without, however, having any function in this mouse. Only after cross-breeding of the donor mouse with a suitable transgenic recipient mouse was the protein to become effective and activate specific genes of the recipient mouse.
The donor mouse was prepared by insertion-mutagenesis during the realization of a transgenic mouse project. The actual goal of the project consisted in establishing transgenic mice which express the tetracycline-controllable transactivator (tTA) in lymphoid cells. The expression cassette used for microinjection into pronuclei comprised the following elements in the 5' to 3' orientation: μE: enhancer from the

intron of the heavy chain of the immunoglobulin genes of the mouse (700 bp); a synthetic promoter consisting of an octam r oligonucleotide and of the minimal promoter of the mouse-ß-globin gene (Wirth et al., Nature 329 (1987), 174-178) and a Tet-R/VP16 construct. The enhancer/promoter combination has been described in Annweiler et al. (Nucl. Acids Res. 20 (1990), 1503-1509). The Tet-R/VP16 construct has been described in Gossen and Bujard (Proc. Natl. Acad. Sci. USA 89 (1992), 5547-5551). The overall size of the DNA fragment is about 3 kb.

In order to prepare the transgenic mice, 1 to 2 picoliters of a DNA solution containing the above-described expression cassette (concentration 1 ng/µl) were injected into the male pronucleus of a fertilized ovum of an NMRI mouse. Subsequently, the ovum was transplanted into the oviduct of a pseudopregnant female foster mouse and was carried by this foster mouse to full term.

Transgenic donor mice normally do not show a phenotype, as the artificial gene is simply injected into the fertilized ovum and integrates in any region of the murine genome purely on a random basis.

As only about 5% of the genome comprise coding regions, the probability that a defect is caused in an essential gene is correspondingly low. Moreover, the mammalian genome is diploid, that is to say all genes are present in duplicate. As a possibly mutated gene, as a rule, has a fully functioning copy as a counterpart which can compensate for the defect in the mutated version, most mutations are recessive, that is to say, they are not expressed if only one copy of the gene is affected.

One of the founder animals obtained during the production of the above-described donor mice now surprisingly showed an extremely conspicuous phenotype in that it was distinctly larger than the siblings of the same litter. The distinctly longer tail and the longer limbs, in particular the long toes were conspicuous. The difference in size compared to normal mice significantly increased in the subsequent weeks and a marked scoliosis formed. All bones except for the scull bones are 1.3 to 1.5 times longer. Consequently, the transgenic mouse is altogether about 1.5 times longer than a corresponding wildtype mouse (see Figure 1). Because of the greatly elongated bones (see Figure 7), the transgenic mouse was termed LOBO mouse (for LOng BOnes). In mice, bone growth comes normally to a standstill in the course

of the development of the individual. In the case of the LOBO mice, the bones of the animal seem to grow incessantly up until the animal's death. In adult animals, this leads to a deformation of the whole individual (see Figure 8) which can be such that the animals can no longer move and female mutants - apart from very few exceptions - can no longer be mated.

The further histological analysis of bones of transgenic mice showed significantly thickened growth zones (see Figure 4). On the one hand, this thickening is attributable to the fact that the number of cells (chondrocytes) is distinctly increased both in the proliferative zone and in the hypertrophic zone. This has been shown not only microscopically, but also immunohistochemically with antibodies against collagen X. On the other hand, the hypertrophic chondrocytes are also larger in the mutants than in the wildtype. Another reason for the increased bone growth resides in the fact that the epiphyseal cartilages (= bone growth zones) in the mutant animals close later than in the wildtype, that is to say, that chrondrocyte proliferation and differentiation proceed longer. At present, it is unclear, whether this proliferation will ever stop completely, as the animals die after about 6 to 8 months for as yet not completely elucidated reasons. Up to said time, the bones seem to continue to grow.

As already mentioned, the mutant animal has a lower life expectancy than its wildtype siblings; about 6 weeks after their birth, LOBO mice show higher mortality, and after almost a year all mice have died for as yet unknown reasons. Homozygous mice are viable. Although so far only two litters of homozygous animals have been obtained, the homozygous animals are born in the expected number. Just as the heterozygous animals they show the increased bone growth which can unambiguously be seen from the longer toes.

Example 2

Genetic analysis of the transgenic mouse

The molecular analysis of the reason for the mutation showed that about 1.5 copies of the transgene were inserted into the intron of an endogenous gene. The insertion is located at 48.2 kb from exon 8 and 5.6 kb from exon 7 (see Figure 9) and has led to the deletion of 11 base pairs. All so far identified exons of the LOBO gene are also present in the transgenic LOBO mice and unchanged vis-à-vis wildtype sequences. Expression studies (Northern analyses) with a cDNA probe of the endogenous gene showed that the gene in question is obviously ubiquitously expressed. While most organs show only one single band (about 4 kb) in Northern blot, the liver shows an additional shorter transcript (about 2 kb). It is unclear whether this smaller transcript a) represents a splice variant of the gene, b) is attributable to the use of an alternative promoter or c) represents the cross reaction with a related gene. Compared to the wildtype animals, only about 50% of mRNA is found for this gene in the heterozygous animals if a probe from the 3'- region of the insertion site is used.

Example 3

Identification and Characterization of the LOBO Gene

In order to find out which gene(s) is/are responsible for the LOBO phenotype, the mutated region from the transgenic mouse was subcloned in bacteria. Localization of the mutated region in the murine genome and subsequent subcloning were possible because the nucleotide sequence of the artificial gene mentioned at the beginning was known and this information could be used in corresponding molecular biological experiments. For the identification of the gene which is called "LOBO gene" hereinafter, 6 kb were sequenced from the subcloned region of the transgenic mouse and at first 87 kb (see SEQ ID Nos. 5 and 6) and then 138 kb (see SEQ ID No. 10, 11 and 12) were sequenced from the corresponding homologous region of the wildtype mouse. The first sequenced region of the murine genomic DNA clone is depicted in SEQ ID Nos. 5 and 6. The sequenced region comprised a total of 86902 base pairs. For technical reasons, this region was divided into two regions, the first 49999 base pairs being depicted in SEQ ID No. 5 and comprising one exon and the remaining 36901 base pairs adjacent to this region at the 3'-end being depicted in SEQ ID No. 6. The exons are localized at the following positions:

SEQ ID No. 5:	8520 - 8753
SEQ ID No. 6:	12487 - 12660
	15497 - 15644
	15908 - 16038
	16148 - 16252
	17293 - 17394
	18083 - 18556

The open reading frame starts at position 8520 in SEQ ID No. 5. The stop codon is located at position 18202 in SEQ ID No, 6. The coding region encodes the amino acid sequence depicted in SEQ ID No. 2. A detailed computer analysis of the first obtained sequence data led to the identification of a gene which consists of at least 8 coding sections ("exons"). The first identified, coding region which is depicted in SEQ ID No. 1 carries the information for 393 amino acids. An overview of the sequenced murine clones obtained in the subsequent sequencing of the 138 kb region is schematically depicted in Figure 10. The sequenced region comprises altogether 138884 base pairs (see SEQ ID Nos. 12 to 15) and contains 12 exons. The exons are localized at the following positions:

Exon	Length [bp]	Start	End
12	80	1117	1196
11	113	30111	30223
10	108	43790	43897
9	234	60504	60737
8	80	91485	91564
7	184	114459	114642
6	87	115272	115358
5	148	117479	117626
4	131	117890	118020
3	105	118130	118234
2	102	119275	119376
1	470	120065	120534

The open reading frame starts at position 1118 in SEQ ID No. 10. The stop codon is located at position 120185.

A detailed computer analysis of the genomic sequence data led to the identification of a gene consisting of at least 13 coding segments ("exons") and being at least 120 kb long, but probably much longer.

The exons identified by genomic sequencing allowed a complete cDNA to be isolated. It is represented in SEQ ID No. 8 and is 3100 bp long. The polyadenylation

signal starts at base 3067, the poly-A tail starts at position 3083. The coding region of the cDNA is 2610 base pairs long. It starts in SEQ ID No. 8 at position 125 and ends at position 2734. The stop codon starts at position 2735. The coding region generates a 870 amino acid long protein, the sequence of which is depicted in SEQ ID No. 9. So far, only the region of position 1243 to position 3083 (start of the poly A tail) of the cDNA in SEQ ID No. 8 has been genomically identified by the 12 exons listed above in tabular form. So far, the cDNA sequence of positions 1 to 1242 has not yet been sequenced genomically, that is to say the intron/exon structure of the gene and its regulatory signals are as yet unknown.

On the basis of the murine sequence data, a DNA probe has been constructed, by means of which a human P1 clone carrying the human LOBO homologous gene, has been isolated. The first obtained sequence of the human genomic clone is depicted in SEQ ID No. 7. The exons are located at the following positions:

1	-	136
3971	-	4118
4500	-	4630
4762	-	4866
5904	-	6005
6600	-	7109

The first nucleotide of the open reading frame is at position 2. The stop codon is located at position 6759. The amino acid sequence represented by the coding region is depicted in SEQ ID No. 4. A clone containing the human genomic sequence was deposited under the accession No. DSM 12073. The first available sequence data showed that the human gene, too, has so far only partially been cloned. An overview of the first obtained and sequenced clones from mice and humans is schematically shown in Figure 2. In order to allow the remainder of the human gene to be sequenced, two further human clones were identified, using the sequence of the human P1 clone, one of said two clones overlapping with the already existing clone in the 5' region and the other in the 3' region. Sequencing of

these altogether 3 clones results in a 311 kb long, human sequence segment depicted in SEQ ID Nos. 15-21. (For technical reasons, the regions have been depicted one after the other with 49,999 base pairs each). The human LOBO exons are localized at the following positions:

Exon	Length [bp]	Start	End
11	113	2701	2813
10	108	13422	13529
9	234	27391	27624
8	80	64694	64773
7	184	94467	94650
6	87	95344	95430
5	148	98485	98632
4	131	99014	99144
3	105	99276	99380
2	102	100418	100519
1	492	101114	101605

The first nucleotide of the open reading frame is located at the genomic position 2703. The stop codon is located at position 101273. The human genomic LOBO sequence contains 4 gaps, each of which is at the most 100 base pairs wide. These gaps are located at the following positions:

Gap 1: 11805 to 11836
Gap 2: 35184 to 35199
Gap 3: 191949 to 191975
Gap 4: 251627 to 251646.

As all sequencing gaps are exclusively located in introns, the coding region remains unaffected. The coding region covered by the exons and the amino acid sequence encoded thereby are depicted in SEQ ID Nos. 13 and 14, respectively. A bacterial clone containing the human genomic sequence has been deposited under DSM

12715. The existing sequence data show that the human LOBO gene, too, has so far only partially been cloned. An overview of the human clones obtained and sequenced is schematically depicted in Figure 10.

Example 4

Chromosomal localization of the LOBO gene

One of the mouse clones obtained which represents a part of the murine LOBO gene was color-labeled by "Fish" (fluorescent in situ hybridization), and hybridized to complete murine (metaphase-) chromosomes. A color signal resulted in band 1D on chromosome 1 of the mouse. This region is homologous to band 2q35-2q37 on human chromosome 2. The result of this experimental mapping is confirmed by the sequence data: The STS marker WI-8964 which is mapped on 2q37 follows 73 kb behind the human LOBO gene. This marker is flanked by 3 phosphatase genes and 2 genes for a nicotine-dependent acetyl choline receptor (see Figure 10). These genes have also been mapped to 2q37 with the result that the chromosomal localization of the human LOBO gene has been unambiguously verified.

Example 5

Expression of the LOBO gene

Expression in the wildtype mouse:

Expression studies (Northern blot analyses) with a cDNA probe of the LOBO gene showed that the gene at issue is ubiquitously expressed. While most organs only produce one single about 4 kb long band in Northern blot, the liver is found to have an additional, shorter transcript (about 2 kb). For the time being, it is still unclear whether this small transcript (a) represents a splice variant of the gene, (b) is

attributable to the use of an alternative promoter, or (c) represents the cross reaction with a related g ne.

Expression in heterozygous and homozygous LOBO mice:

In Northern blot only about 50% of the LOBO mRNA is found in heterozygous mice compared to the wildtype, while no LOBO mRNA can any longer be detected in homozygous mice. Hence, the artificial DNA insertion can be assumed to produce a disorder in the maturation of the mRNA. In this process, the introns which are still contained in the primary RNA are cut out (splicing). This cutting out is brought about by certain sequence signals. Such signals are also contained in the artificially inserted gene, with the effect that presumably a so-called aberrant splicing occurs. As a consequence, a functioning LOBO mRNA is prevented from being formed, and the corresponding protein cannot be produced, at least not in its full length. As the transcription signals of the LOBO gene are not affected by the insertion of the transgene, at least a shortened and moreover chimeric LOBO mRNA could be expected to be produced from the natural transcription start to the splice signal in the inserted sequence. However, a polyadenylation signal is missing in the transgene-insertion, which leads to a non-polyadenylated RNA which should show a distinctly lower stability than the normal mRNA. That is to say, the amount of this chimeric RNA should be rather small and below the Northern blot detection limit. In fact, this chimeric RNA has not been detected in Northern blot so far. However, with the much more sensitive RT-PCR method it has been possible to verify the existence of this postulated chimeric RNA. It can be assumed that this RNA prompts the formation of a shortened LOBO protein which possibly also performs partial functions of the complete LOBO proteins or competes with it for binding partners or for the substrate.

Expression in human tumor tissues:

The sequence of the LOBO protein derived from human cDNA shows high homology to the human Dis3-gene. For this gene, a Japanese working group has shown that its expression rate in tumor tissues was distinctly altered compared to the corresponding normal tissues. In order to examine whether the LOBO gene

behaves analogously, a commercially available Northern blot which was charged with RNAs from different tumor tissues was hybridized to a human LOBO probe. The different tumor types in fact showed significant expression differences (Figure 5). However, the biological interpretation of these data is difficult. Nevertheless, the LOBO gene might possibly play a part in carcinogenesis.

Example 6

Characterization of the LOBO protein

The murine and human amino acid sequences derived from the LOBO cDNAs were compared with known proteins. This comparison showed that the amino acid sequence has regions highly conserved between organisms ranging from mammals (mouse and humans), to invertebrates (Caenorhabditis elegans), unicellular eukaryotes (Saccharomyes cerevisiae, Schizosaccharomyces pombe) and prokaryotes. A relationship analysis of these proteins shows that the murine and human LOBO proteins represent a group of their own (see Figure 6) which is, however, related to two other protein groups. One group comprises the VacB and the RNAse type II proteins from bacteria, the VacB proteins having been found to also possess type II RNAse activity, according to a recent publication. A second group comprises the Dis3-homologous proteins from different eukaryotes ranging from mammals to unicellular yeasts.

The clear relationship to the two afore-mentioned protein groups makes it possible for the function of the LOBO proteins to be estimated, as the LOBO proteins can be assumed to also have similar functions because of their structural similarity to the afore-mentioned groups of protein. On this basis, the following functions can be postulated for the LOBO protein:

(a) it plays an important role in the cell cycle regulation (mitosis control) (proven for Dis3 from S. pombe; here, the gene's loss of function leads to the loss of the cell's capability to divide);

(b) because of its bearing on the cell cycle control, the conclusion suggests itself that the LOBO protein possibly also plays a part in carcinogenesis (proven for Dis3 from Homo sapiens; the results depicted in Figure 5 support the above-mentioned assumption).

(c) The LOBO protein most probably has the ability to bind RNA (proven for the LOBO-type SSDI protein from S. cerevisiae and for the VacB and RNAse type II proteins).

(d) The LOBO protein has at least one protein binding partner. It is presumably a G-protein or a G-protein-controlling protein (proven for Dis3 from S. pombe which binds to the G-protein-regulator RCC1 and controls its activity).

Example 7

Clinical relevance of the human LOBO protein

Sequencing of a genetic STS marker (WI-8964) in the 3' region of the LOBO gene has made its chromosomal localization in humans known. The human LOBO gene is positioned on chromosome 2, band q37. In this region, a hereditary disease has been mapped which leads to a bone growth disorder in humans, the so-called "Albright hereditary Osteodystrophy" (AHO). AHO is a syndrome consisting of a number of different symptoms pronounced in varying degrees, depending on the patient. However, three of these symptoms are characteristic of this disease and appear in all patients: hyposomia, obesity, brachydactylia. It is known from the literature that this disease is mapped on two different sites at the same time: at the above-mentioned position (2q37) and moreover on chromosome 20, band q13. The gene on 20q13 responsible for AHO is a G protein, the loss of function of which leads t- the typical AHO symptoms. However, there are also AHO patients, who do

not show any defect in respect of 20q13, but show a def ct (mostly a deletion) in 2q37, and nevertheless show the AHO phenotype. It is therefore possible that two proteins, one of 20q13 and one of 2q37, directly or indirectly interact and jointly perform a function. In the case of a defect in one of the two protein partners a loss of function or malfunction would occur and possibly cause a visible phenotype. As the gene of 20q13 is a G-protein and LOBO stems from 2q37, and moreover has a great similarity to (Dis3) proteins, which indirectly control G-proteins, the conclusion suggests itself that LOBO is the candidate gene for "Albright hereditary osteodystrophy". The fact that AHO patients suffer from hyposomia, while LOBO mice show exaggerated growth may be attributable to the type of mutation. The type of mutation which is present in the mouse (insertion of an artificial gene) is artificial, and certainly is not found in AHO patients. In this case, large deletions which are likely to delete the whole LOBO gene are the prevalent mutation type. An example where a gene can cause both hyposomia and megasomia, depending on the type of mutation, has been published. Moreover, the same mutation of one and the same gene in a mouse or in a human can lead to quite different phenotypes, because these organisms are different in many respects.